

Agreement Number: 03HQAG0177

Category 3: Don't Duck Metadata – Metadata Outreach

Final report

Organization: Barrow Arctic Science Consortium

Glenn Sheehan, Executive Director, Glenn.Sheehan@arcticscience.org,

POB 577, Barrow Alaska 99723 www.arcticscience.org

Project Leader: Allison Graves Gaylord, nunatech@usa.net

Nuna Technologies PO Box 1483, Homer, AK 99603, USA. Tel / Fax : 907-235-3476

Collaborating Organizations: Name, point of contact, address, web address

NSF Arctic System Science (ARCSS) Data Coordination Center (ADCC)

Rudolph Dichtl, Manager

The National Snow and Ice Data Center, University of Colorado, Campus Box 449,

Boulder, CO USA 80309-0449

University of Texas El Paso

Dr. Craig. E. Tweedie, tweedie@msu.edu

500 W. University Ave, El Paso, TX 79968

Arctic Ecology Laboratory

Dr. Patrick Webber, webber@pilot.msu.edu

100 North Kedzie Hall, Michigan State University, MI 48824-1031

National Geophysical Data Center, NOAA

Nancy Auerbach, nancy.auerbach@noaa.gov

Mail Code E/GC, 325 Broadway, Boulder, CO 80305-3328.

Cooperative Institute for Research in Environmental Sciences

Leanne Lestak, leanne.Lestak@colorado.edu

University of Colorado, Boulder, CO 80309-0216

North Slope Borough Department of Planning and Community Services

Paul McNeil, paul.mcneil@north-slope.org

P.O. Box 69, Barrow AK 99603

Geographic Information Network of Alaska

Dr. Virgil Sharpton, buck.sharpton@gi.alaska.edu

Geophysical Institute, University of Alaska Fairbanks, 903 Koyukuk Dr.,

P.O. Box 757320 Fairbanks, AK 99775-7320

International Permafrost Association

Jerry Brown, jerrybrown@igc.org

PO Box 7, Woods Hole, MA 02543

Institute of Arctic and Alpine Research
Dr. William Manley, william.manley@colorado.edu
University of Colorado, Boulder, CO 80309-0450 USA

Project Narrative

Summarize the project activities. Include its accomplishments, successes, strengths and weaknesses, further challenges, and collaboration activities as appropriate. What are the organizational practices that have developed to support metadata creation and maintenance for the future? In writing the report keep in mind the goals of your project under this category: the development of new or strengthening of existing multi-organizational collaboration that supports the development and maintenance of shared digital geographic resources, and to foster the establishment of cross-organizational coordinating councils that develop and advance the NSDI within a specific geographic area.

The USGS Metadata Outreach Agreement with the Barrow Arctic Science Consortium has furthered the development of best practices among the research community to increase access to spatial data for the region surrounding Barrow, Alaska through an effort known as the Barrow Area Spatial Data Infrastructure. Metadata development is fundamental to this multifaceted effort that included matching funds from the National Science Foundation (NSF) to physically document current and extant research sites in the region with Differential GPS technology, to make this information available via the Barrow Area Information Database – Internet Map Server (BAID-IMS) prototype and to archive data at the newly established National Spatial Data Infrastructure (NSDI) clearinghouse node at the Arctic System Science (ARCSS) Data Coordination Center (ADCC). In addition to the NSF, BASC partners include NASA, NOAA, DOE, EPA, USFWS, the North Slope Borough and Ukpeagvik Iñupiat Corporation (UIC).

BASC facilitates about 40 research projects that represent the needs of numerous universities and partners with interests in the acquisition and use of spatial data in the Barrow area. These needs range from the permitting of new research activities to comparative analysis of legacy data associated with environmental changes over time.

As project lead, Allison Graves Gaylord has been intimately involved in both the technical aspects of metadata development and outreach activities to strengthen multi-organizational collaboration in support of the development and maintenance of shared digital geographic resources associated with the Barrow Area Spatial Data Infrastructure. Technical aspects include conducting hands on training, XML template development, metadata compilation, quality control, parsing, archiving activities with the ADCC clearinghouse, integration of metadata with the Barrow Area Information Database – Internet Map Server and development of a recommended metadata protocol. Outreach activities include meetings, presentations and interviews with scientists, government contacts and native corporations plus the development of outreach materials including

web pages (<http://ims.arcticsscience.org/dwg/>), posters (see appendix A and B) and associated articles (see appendix C.)

Metadata outreach included Powerpoint presentations and live demonstrations to the following groups: Alaska Ocean Observation System Data Management and Communications (DMAC) Committee, Arctic Coastal Dynamics Working Group, BASC's Science Advisory Group, BASC's Science Management Advisory Committee, Barrow School Yard, plus an evening presentation to the Barrow community and a talk hosted by the University of Alaska Fairbanks Geophysical Institute, as well as two presentations at the 2003 Study of Environmental Arctic Change (SEARCH) workshop held in Seattle, WA.

Challenges

- Staff turnover at key institutions during the project was a challenge but re-enforced the need for metadata development as a means of preventing the loss of institutional knowledge (including the resignation of the technical lead at the ADCC, personnel turnover at the NSB GIS and UAF, plus graduating students).
- While there are incentives among the scientific community to publish papers, the same incentives do not exist to publish research data. There is a need to change this "culture" to place equal value on publishing data.
- There is a need to encouraging agencies, contracting entities and principle investigators to budget time for metadata development and data archiving activities for every project or contract. Contracting entities and agencies should clearly spell out requirements to develop FGDC metadata and contribute final data sets to the national clearinghouse system or Geospatial One-Stop portal.
- Numerous XML based metadata templates have been developed for various projects, and while there has been ongoing support available through this outreach grant, some entities have not made the time to complete metadata records. This is largely due to a lack of metadata being a requirement of their own contract and grant requirements.
- Parsing metadata for the clearinghouse can be a challenge for several reasons: (1) Iñupiaq characters (2) Differences between versions of the USGS Metadata Parser (MP) are commonly in use. ESRI's ArcCatalog software uses an older version of the parser that produces erroneous errors. It was determined that metadata records should be parsed outside of ArcCatalog with the DOS based tool (MP Version 2.8.7 or higher) or USGS's online validation tool before transferring the files to a clearinghouse node. (3) Matching MP errors with the required entries in ArcCatalog can be elusive. (4) Parsing is very time consuming and often results in the need for extensive follow up with data producers.
- Low bandwidth Internet connections (outside of the BASC facilities) at Barrow hampered informal outreach efforts among North Slope Borough staff members, Alaska Department of Fish and Game Biologists and Ukpeagvik Iñupiat Corporation staff.
- Community presentations at the Ukpeagvik Iñupiat Corporation Science Division and BASC facilities were somewhat limited due to space. Hands on demonstrations were conducted with student assistance at BASC at computer

terminals distributed through labs and the science library. The Barrow High School GIS lab was being moved to a new building during planned outreach sessions, but would be the ideal venue for training.

- The lack of remote access to a server at Barrow resulted in utilizing a server at the Arctic Ecology Lab at Michigan State University.
- While not part of this grant request, the synchronization of metadata records between the ArcIMS prototype (BAID-IMS) and the NSDI clearinghouse is a challenge.
- The ADCC Clearinghouse moved web harvestable folders for this project twice which resulted in the need for the project lead to revise metadata templates, update existing metadata records and metadata records published in the BAID-IMS prototype web portal.

Measurable Project Results:

- *Indicate how metadata is served or posted*

Over 50 metadata records are linked to data layers in the Barrow Area Information Database-Internet Map Server (BAID-IMS) prototype with a special focus on finalizing metadata associated with the Science Research District. Metadata records associated with the remaining layers in BAID-IMS have been drafted and are undergoing quality assessment / quality control. BAID-IMS can be accessed at: <http://www.baidims.org> The National Science Foundation has just funded BAID-IMS as a 3 year project.

Due to agency contracting requirements for archiving, more than one NSDI clearinghouse node is being utilized. Data developed for NSF funded activities is primarily housed at the Arctic System Science Data Coordination Center (ADCC) at the National Snow and Ice Data Center (NSIDC) at the following links: <http://nsidc.org/arcss/> and <http://nsidc.org/data/gis/data.html>

A second regional archive utilized for research data is the Geographic Information Network of Alaska housed at the University of Alaska: <http://map.gina.alaska.edu/metadataexplorer/explorer.jsp> Hundreds of image mosaics and derived sea ice data (in ArcGIS Grid and Shapefile format) will be available at the following clearinghouse node in the December, 2005 at the completion of associated MMS contracts: <http://map.gina.alaska.edu/metadataexplorer/explorer.jsp>

Examples of web harvestable folders include project folders listed at:

<ftp://sidads.colorado.edu/pub/DATASETS/ARCSS/data/>

<ftp://sidads.colorado.edu/pub/DATASETS/ARCSS/data/arcss400/>

<ftp://sidads.colorado.edu/pub/DATASETS/ARCSS/data/arcss031/>

<ftp://sidads.colorado.edu/pub/DATASETS/ARCSS/data/arcss301/>

<ftp://sidads.colorado.edu/pub/DATASETS/ARCSS/data/arcss303/>

- *Indicate how many metadata entries were created*

Over a 200 metadata records have been drafted, many of these are still undergoing quality control and parsing to ensure they will easily ingest into the clearinghouse system. In several cases metadata development was initiated at the beginning of a multi-year study and the principle investigators did not want to “publish” the data until the end of the project. Completion of these remaining records is anticipated during 2005-07 with funding provided by that National Science Foundation.

- *Do you need assistance in providing for metadata service to organizations you have assisted?*

Two utilities would be helpful. (1) A free HTML to XML conversion tool would be very useful for updating existing metadata records. This issue came up often and it seemed silly to re-type information into ESRI's ArcCatalog software in order to have an FGDC compliant file that was associated with a data set. I'm also interested in learning more about metadata “crosswalk” tools for updating metadata records that were originally compiled in non-FGDC formats (e.g. ISO, DIF, EML, SensorML, MarineXML, Dublin Core, etc.)

While this was not part of this grant request, training is desired in the synchronization of the BAID-IMS application with the NSDI clearinghouse system. Currently, BAID-IMS has been a prototype and NSF funds have been requested to fully develop the application.

Metadata training and outreach assistance

- *List organizations and number of individuals receiving metadata training and outreach assistance as appropriate*

Over a dozen long-term Barrow area scientists were interviewed in an effort to compile metadata records associated with major research initiatives in the region dating back to the establishment of the Naval Arctic Research Laboratory in 1947. Extensive follow up via email included the review of html formatted metadata records. In some cases, a basic Microsoft Word metadata template was devised to make it easy to populate legacy metadata records. Ten young investigators from the Arctic Ecology Lab at Michigan State University, the university of Alaska Fairbanks, the University of Cincinnati and the University of Delaware received hands on technical training in the use of metadata templates, parsing and data discovery through the NSDI clearinghouse system. Four staff members of the North Slope Borough GIS Division received training and ongoing support to develop metadata records for data associated with the North Slope villages and subsistence information

compiled for maps published in the *Guidelines for Improved Cooperation between Arctic Researchers and Northern Communities*. The use of the FGDC standard was demonstrated to Differential GPS specialists from UNAVCO (<http://www.unavco.org>) and UNAVCO staff provided assistance in compiling a Differential GPS template for Barrow researchers. In addition, both formal and informal outreach was conducted with local agencies and native corporations including the North Slope Borough Planning Director and land managers with the Ukpéagvik Iñupiat Corporation and Arctic Slope Regional Corporation.

The project lead participated in the Arctic Coastal Dynamics (ACD) workshops in the Fall of 2003 and 2004. Over 50 international participants attended each of these workshops. Poster presentations were accompanied by demonstrations of the FGDC standard, NSDI clearinghouse system and BAID-IMS (see appendix A and B.) In addition, the project lead is an active participant in the ACD GIS Working Group (10 participants) and has encouraged the adoption of the FGDC/ISO standard for archiving activities. Allison Gaylord is also a co-author of the GIS chapter of the ACD book in which SDI and metadata activities are being promoted internationally (draft in progress.) See associated articles appendix C.

Demonstrations of the use of metadata and data discovery tools were conducted at the Alaska Ocean Observation System Data Management and Communications Committee workshop in March, 2004 and January, 2005. The FGDC standard and FGDC profiles was recommended to AOOS DMAC. Portal technologies such as Internet Map Servers (including BAID-IMS, the US National Map, etc.) and the NSDI clearinghouse house were demonstrated to an audience over 20 participants at these workshops.

- *At what level of proficiency are the trainees or training (introductory to advanced)*

Training activities varied depending on the technical proficiency of the audience. Hands on technical instruction on the use of metadata templates in ESRI's ArcCatalog software occurred in Barrow over the course of the project (for a total of at least six weeks on site technical support.) This training was primarily targeted at young investigators and the North Slope GIS.

- *Indicate the number and character of workshops conducted as appropriate*

Outreach presentations included talks, demonstrations and posters at the following venues attended by Barrow area researchers:

1. Alaska Ocean Observation System Data Management and Communications (DMAC) Committee, Anchorage AK (2004, 2005)
2. Arctic Coastal Dynamics Working Group, St Petersburg, Russia and Montreal, Canada (2003, 2004)
3. BASC's Science Advisory Group, Barrow, AK (2003)

4. BASC's Science Management Advisory Committee, Barrow, AK (2003)
5. Barrow School Yard, Barrow, AK (2003)
6. University of Alaska Fairbanks Geophysical Institute, Fairbanks, AK (2003)
7. Presentations at the (2003) Study of Environmental Arctic Change (SEARCH) workshop held in Seattle, WA.
8. IEEE International Geoscience and Remote Sensing Symposium held in Anchorage, Alaska (2004)
9. American Geophysical Union, San Francisco, CA (2004)
10. BASC Science Talk for the Barrow Community, Barrow AK (2005)
11. MMS Information Transfer Meeting, Anchorage, AK (2005)

Next Steps

- *Will this project's activities continue in the future?*

The development of this project is the direct result of needs identified by the Digital Subcommittee of BASC's Science Management Advisory Committee (SMAC) in 2003. During the BASC meetings held in conjunction with the American Geophysical Union conference in late 2004, it was recommended that metadata outreach activities continue to be promoted through the SMAC Digital Subcommittee.

During the annual meeting of BASC's Science Management Advisory Committee, it was discussed that the metadata development, data archival and online interface development associated with the BAID-IMS prototype had intellectual merit and were worthy of an independent proposal to NSF (independent of the logistic support activities covered by the BASC-NSF cooperative agreement.) Two pending proposals were submitted to the National Science Foundation on August 30, 2004 that are directly related to this effort: (1) *Maintenance, development and innovation of the Barrow Area Information Database and Internet Map Server (BAID-IMS)* (2) *Collaborative Research: Developing and Implementing an Arctic Spatial Data Infrastructure to Support International Arctic Science*. The first proposal has been funded and the second is still pending reviews. Both efforts include ongoing metadata activities, data archival and the implementation of Geospatial One-Stop technologies.

In another related international effort, the need to continue targeted outreach about best practices for handling spatial data among the scientific community is also the focus of an Arctic Spatial Data Infrastructure proposal submitted by the project lead as an International Polar Year Expression of Intent.

The project lead is also serving on the Alaska Ocean Observation System (AOOS) Data Management and Communications Committee (DMAC). AOOS DMAC is developing standards for handling data associated with the emerging regional observatories in Alaska. Allison Gaylord has given two presentations on the use of

IMS portal technologies (including a demonstration of BAID-IMS, the US National Map and clearinghouse search and example FGDC metadata records.)

A synergistic activity related to this award is the capturing of grey literature associated with historical research in the Barrow area, which is being spearheaded by the advisory groups of the Barrow Arctic Science Consortium. Through this ongoing data rescue effort, support has been provided by the Geophysical Institute Library at the University of Alaska Fairbanks where a collection of donated grey literature is being compiled from the personal archives of retired Barrow area researchers. The collection includes dissertations, theses, government reports and old maps of the Barrow region that are not widely published. This effort is also supported by the National Science Foundation.

- *Describe the next phase in your project*

- Coordination with the Department of Interior's North Slope Science Initiative Data Management Group will be initiated in April, 2005.
- Coordination with the Barrow native corporation contractor regarding easement and trail mapping associated with the Alaska Native Claims Settlement Act.
- Follow up training and assistance was conducted with the NSB GIS staff in March, 2005. Ongoing on site follow up is planned in May and August, 2005 (in addition to teleconferences and email follow up.)
- Follow up interviews with long term North Slope researchers meeting at Barrow are planned for May and August 2005.
- The project lead will be available at Barrow in May and August to provide training and assistance to a new pool of young scientists working at Barrow.
- Ongoing quality control for metadata records under development will be completed.
- Additional parsed metadata records will be submitted to the clearinghouse.
- An XML template under development for the Arctic Coastal Dynamics program is undergoing final review.
- XML template development for the Alaska Ocean Observation System is planned.
- A protocol for submitting sensitive cultural information to the archive will be discussed with the Inupiaq History Language and Culture Commission.
- A Geospatial One-Stop harvestable web folder will be created in conjunction with ongoing development of BAID-IMS.

- *Requirements (more technical assistance, software, other?)*

Instruction on the synchronization of existing IMS applications (such as BAID-IMS) with the national Geospatial One-Stop system for seamless metadata harvesting is desired.

- *What areas need work?*

- Additional follow up to complete draft metadata records. Better synchronization with the NSDI.

Feedback on Cooperative Agreements Program

- *What are the program strengths and weaknesses?*

The FGDC / NSDI web site is a tremendous resource, but it's not always easy to find information. Metadata outreach was particularly successful in cases when it is clear that metadata is a required deliverable for a grant or contract. Many of the scientists and other government funded contacts I worked with were surprised to hear that metadata is a requirement for many federal initiatives. Metadata and data archiving requirements should be listed clearly on federal grant announcements and under the deliverables section of all federal RFPs and contract documents. State and local agencies, as well as non-profit entities contracting the development of spatial data should be made aware of this shortcoming and also clearly state that FGDC metadata should be a deliverable.

During this outreach, several questions were raised about other non-FGDC metadata activities that are funded by federal monies. For instance, the National Science Foundation has funded research on the use of the Dublin Core standard. While NOAA Coastal Services Center conducts excellent outreach on FGDC metadata (and lists this requirement on their contract documents), NOAA is also funding the Integrated Ocean Observation System (IOOS) and regional IOOS efforts are considering a variety of metadata tools for marine data (SensorML, MarineXML, DIF, Dublin Core, etc.) It should be widely promoted that FGDC metadata is not just for spatial data and that profiles can be used to extend the standard to address the needs of specific domains. Allison Gaylord conducted presentations to the Alaska Ocean Observation System Data Management and Communications Committee about the flexibility of the Content Standard for Digital Geospatial Metadata (CSDGM).

FGDC CAP award contacts should be informed about the significance of the International Organization for Standardization (ISO) Technical Committee's [(TC)211] Metadata Standard 19115 that was adopted June 2004 and how this may apply to the utilization of FGDC metadata tools. Many BASC scientists who conduct work in Alaska also have similar projects in other Arctic nations. Questions about the difference between federal and international metadata standards are often raised. The differences are not clear.

- *Where does the program make a difference?*

Outreach activities were particularly successful among young investigators and retired/semi-retired scientists. Young investigators are increasingly engaged in using digital data as part of their research and often struggle to find data for use in thesis and dissertation projects. These young investigators were most impressed by the ability to query the clearinghouse system for data sets they helped create. Many long-term Barrow researchers are eager to see their work carried on and understand the value of properly archiving.

- *Was the assistance you received sufficient or effective?*

I received an announcement about FGDC Train the Trainer workshops, but the timing for these sessions was not convenient. Peter Schweitzer was very helpful and responsive via email. In addition, I received support from ESRI regarding parsing issues pertaining to metadata records composed in ArcCatalog 8.3. The National Park Service metadata tools extension for ArcCatalog proved to be a tremendous resource for generating templates and applying them to directories of similar data sets.

- *What would you recommend doing differently?*

-Hold an annual workshop to kick off the FGDC CAP awards. This could include a few key presentations from efforts funded the previous year (or even just posters about those efforts.) In addition, this would be an ideal opportunity to hold a “Train the Trainer” workshop and to provide information about the latest developments regarding the National Geospatial One-Stop effort, NSDI, evolving metadata standards to accommodate ISO, metadata cross walking tools, etc.

-Provide a list of FGDC personnel contacts that are available to answer technical questions.

-All federal funding agencies should require that a minimum of 10% of a research budget is dedicated to data management activities geared toward developing FGDC metadata and contributing final data sets to the National Geospatial One-Stop portal.

- *Are there factors that are missing or need to consider that were missed?*

No.

- *Are there program management concerns that need to be addressed? Time frame?*

The one-year time frame for the project was not enough, so a three-month extension was requested. Although meetings were initiated with the Arctic System Science Data Coordination Center at the National Snow (ADCC) and Ice Data Center in Boulder, CO shortly after the grant award was announced (in October, 2003), it was discovered that the existing FGDC node at the National Snow and Ice Data Center could not be used for this project. So, the ADCC established a new FGDC node for Barrow area spatial data. The technical point contact at the ADCC resigned in 2004 and the position was not filled. ADCC staffing constraints have delayed ingesting Barrow area data sets in to the clearinghouse.

The BASC Science Management Advisory Committee feels that metadata outreach should be an ongoing activity. It is difficult to constrain the project goals to a single year since the active research community is somewhat dynamic.

- *If you were to do this again, what would you do differently?*

Request that the project time frame be extended through 2005.

Appendices

Appendix A – The Importance of Metadata for Archiving and Promoting Spatial Data Discovery

Appendix B – Barrow Area Information Database – Internet Map Server

Appendix C – Associated Articles

Graves Gaylord, A., et al., The Importance of Metadata for Archiving and Promoting Spatial Data Discovery, ACD-Arctic Coastal Dynamics, 5th International Workshop, October 13-16, 2004, McGill University, Montreal, Canada, (abstract); in press.

Graves, A., Tweedie, C., Zaks, D., Serbin, S., Manley, W., Brown, J., Bulger R., (2004). Spatial Data Infrastructure Supports Long Term Measurements to Detect Arctic Change. Arctic Coastal Dynamics, Report of the 4th International Workshop. St. Petersburg, Russia, 10-13 November 2003. Reports on Polar and Marine Research 482, pp. 52-56.